



Das Computing Grid für den europäischen Teilchenbeschleuniger des CERN

Grenzdimensionen

Bernd Schöne

In der Nähe von Genf betreiben Physiker den größten Teilchenbeschleuniger der Erde, den Large Hadron Collider. Damit das riesige Physiklabor gleichzeitig Tausenden von Wissenschaftlern in aller Welt zur Verfügung steht, hat das CERN zusammen mit anderen Organisationen das derzeit größte Rechen- und Speicher-Grid aus der Taufe gehoben. Nun haben sie das nächste Ziel im Visier: die Brücke vom Grid zur Cloud.

Einer der wenigen Orte, an denen man Informationstechnik der anderen Dimension bewundern kann, ist das CERN, das Europäische Zentrum für Teilchenphysik (Conseil Européen pour la Recherche Nucléaire). Dort bereitet sich das hauseigene Rechenzentrum auf gewaltige Rechenaufgaben und Datenmengen vor. Sie sollen anfallen, wenn der weltgrößte Teilchenbeschleuniger, der Large Hadron Collider (LHC), am CERN im Jahr 2013 seine maximale Leistungsfähigkeit erreicht.

Zwar muss etwa auch der amerikanische Nachrichtendienst NSA (National Security Agency) gigantische Bit-Mengen speichern, um all die mitgeschnittenen Telefon- und Internetverbindungen zu konservieren, doch die Tore dieser Institution bleiben für Besucher leider genauso versperrt wie die Rechenzentren der Atomenergiebehörde, auf deren Großrechnern das US-Militär Atomexplosionen simuliert.

Dagegen macht das CERN aus seiner Tätigkeit kein Geheimnis; das insti-

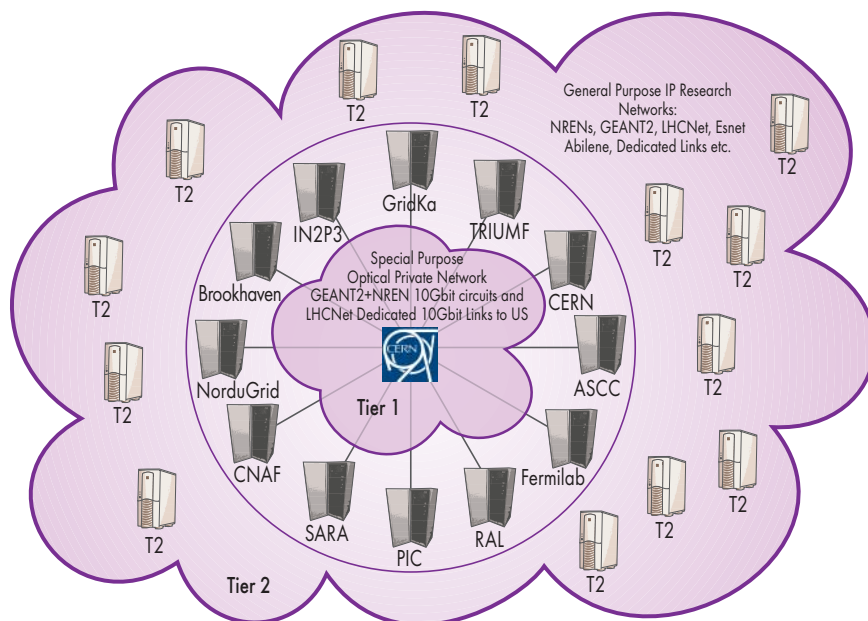
tutseigene Besucherzentrum ist für jeden zugänglich. Ausgewählte Gruppen dürfen sogar das Rechenzentrum betreten, einen Ort, der schon vielen Systemen eine Heimat geboten hat: Mainframes, Vektorrechnern und Unix-Servern jeglicher Art. An deren Stelle entstand einer der größten Linux-Cluster der Welt. Er ist der Ausgangspunkt für ein fünfstufiges weltweit verteiltes Speicher- und Rechen-Grid, das LHC Computing Grid (LCG).

Um die kleinsten Partikel der Materie zu erforschen, hat das CERN einen knapp 27 km langen Ringbeschleuniger plus kleinere Ringe zur Vorbeschleunigung in die Genfer Erde gebuddelt. Genauer gesagt unter Gebiete in Frankreich und der Schweiz. Um im stabilen Gestein arbeiten zu können, liegt der große Ring leicht schief. Unter dem Genfer See verläuft er in 50 Meter Tiefe, in Frankreich sind es 150 Meter. Bevor der Eurotunnel zwischen Frankreich und Großbritannien gebaut wurde, war der Ring das größte Tiefbauprojekt der Geschichte. Je nachdem, wie man rechnet, hat das Ganze mindestens 1 bis 2,5 Milliarden Euro gekostet, viele Sachspenden und die Arbeitsleistung der Wissenschaftler in den Partnerinstituten nicht mitgerechnet.

Als Lohn für die Mühen sollen spektakuläre Erkenntnisse über die Zusammensetzung der Materie und der eine oder andere Nobelpreis winken. Forschungsziel ist es auch, durch die Kollisionen ganz kurzfristig das eine oder andere Teilchen zu erzeugen, das es im Universum seit dem Urknall nicht mehr gegeben hat. Das brachte dem LHC die Bezeichnung „Urknallmaschine“ ein. Die Physiker am CERN hören dieses Wort höchst ungern, haben aber noch keine griffige Alternative.

Keine Chance, alles aufzuzeichnen

Ähnlich kolossal wie die Dimensionen des Experiments mit seinen vier Detektoren ALICE, ATLAS, CMS und LHCb – der größte so hoch wie ein achtstöckiges Haus – sind die dabei anfallenden Datenmengen (siehe Kasten „Der Large Hadron Collider“). Nach den bisherigen Kalkulationen der Forscher wird der LHC 40 Millionen Ereignisse pro Sekunde (40 MHz) erzeugen, bei jedem Zusammenstoß entstehen im Schnitt 100 bis 1000 neue Teilchen, die in alle Himmelsrichtungen auseinanderfliegen. Möglichst viele, so hoffen die Forscher,



Das CERN ist über mehrfache 10-Gbit-Leitungen mit dem Tier-1-Knoten verbunden, diese verteilen die Daten an die Tier-2-Zentren (Abb. 1).

verfangen sich in den Sensoren und hinterlassen dort Informationen über Impuls, Energie und Richtung. Daraus entsteht ein Datenstrom von einer Million GByte/s oder 1 PByte/s.

Beim gegenwärtigen Stand der Technik wäre es vollkommen illusorisch, die Rechen- und Speichersysteme direkt mit den Messdaten der vier Großexperimente zu füttern. Daher gibt es bereits innerhalb der Experimente spezielle Embedded-Rechner, die in drei Stufen eine erste Auswahl der Ereignisse treffen. Die von den Rechnern für relevant erachteten Daten von 1500 Ereignissen pro Sekunde gehen mit 2 GByte/s vom Ring in das ein paar Kilometer entfernte Rechenzentrum. Die Auswahl ist flexibel und geschieht nicht ganz freiwillig. Wenn sie könnten, würden die Physiker gern mehr Daten aufzeichnen, denn umso schärfer ist später das mathematisch rekonstruierte Bild der Vorgänge in der

Kollisionskammer. Selbst für die Auswertung der übrig bleibenden Daten ist das Genfer Rechenzentrum trotz seiner großzügigen Dimensionen zwei Nummern zu klein. Eine Analyse und Ereignissimulation vor Ort würde 100 000 Computer erfordern – angesichts des zur Verfügung stehenden Budgets ein illusorischer Gedanke; schon den Ring und seine Experimente mussten die Genfer auf Pump bauen.

Unerwartete Schwierigkeiten mit wasserführenden geologischen Schichten sprengten den ursprünglich angesetzten Finanzrahmen und zwangen das Management, einen Kredit bei der Europäischen Bank aufzunehmen. Wenn man sich die nötige Rechenleistung nicht selber leisten kann, muss man sie sich eben borgen, dachten die Verantwortlichen.

Ein weltweit verteiltes Rechen- und Speicher-Grid, das Large Hadron Colli-



- 2013 soll der am CERN gebaute Teilchenbeschleuniger Large Hadron Collider die volle Leistung erreichen, die heute schon hohe Datenrate steigt dann weiter. Um die Datenflut zu analysieren, arbeiten nationale Forschungszentren in der ganzen Welt zusammen.
- Eigens dafür hat das CERN mit Partnerorganisationen ein Grid ins Leben gerufen, das die Datenmengen und Rechenanforderungen bewältigen kann.
- Dazu sind in 5 Ebenen über 1000 Institute über eigens dafür entwickelte Netze zusammengeschaltet.

Der Large Hadron Collider: Dem Higgs-Teilchen auf der Spur

Schon immer wollte der Mensch wissen, was die Welt im Innersten zusammenhält. Heute wissen wir: Vier Grundkräfte sind es. Gravitation, elektromagnetische Kraft sowie die starke und schwache Kernkraft. Sie wirken auf Elektronen, Protonen und Neutronen, die unsere sichtbare Welt in Form von Atomen formen. Protonen und Neutronen sind aber nicht wirklich elementar. Sie sind aus jeweils drei Quarks zusammengebaut. Eine auch von den meisten Physikern kaum verstandene Theorie besagt, dass die vier Grundkräfte durch ganze Schauer von Partikeln übertragen werden, die ständig zwischen den Quarks und den Elektronen hin- und herwabern und Kräfte übertragen. Sie sind ein Produkt der Quantenwelt, entstehen aus dem Nichts, verschwinden wieder und existieren nur so kurze Zeit, dass sie den Energieerhaltungssatz der Physik nicht stören.

Der italienische Nobelpreisträger Enrico Fermi legte schon 1934 Beweise vor, dass insbesondere die schwache Kernkraft noch eine weitere Eigenschaft hat: Sie ist indirekt für den radioaktiven Zerfall von Atomkernen verantwortlich. Leider gibt es für eines der mitwirkenden Austauschteilchen außer einem inzwischen fast zum Mythos gewordenen Namen wenig konkretes Wissen: das Higgs-Boson. Es muss vorhanden sein, wenn sich die Kernphysik der letzten dreißig Jahre nicht sinnlos im Kreis bewegt haben soll. Beobachtet hat man es aber noch nie.

Daher sprechen viele Physiker auch lieber von Higgs-Teilchen, schon der Name „Boson“, nach dem indischen Physiker Bose, ist eine Festlegung auf ein Objekt mit ganzzahligem Eigendrehimpuls (Spin). Aber auch die sind ebenso wenig bestätigt wie andere Eigenschaften. Dabei wäre vor allem die Masse des Teilchens nicht nur für die Kernphysik von großer Bedeutung. Denn in allen Sternen laufen permanent Zerfallsprozesse ab, an denen sie beteiligt sein müssen. „Ohne Kenntnis des Higgs keine exakte Theorie der Sonne“ lautet das einfache Resümee. Manche Theoretiker gehen noch weiter, sie glauben, dass Materie überhaupt nur indirekt über eine Masse verfügt. Das Meer von Higgs-Bosonen, das sie umgibt, würde ihnen die Eigenschaft „Masse“ überhaupt erst verleihen. Ganz besonders fatal wäre es, wenn es das Higgs gar nicht geben würde. Dann wären etliche Physiklehrbücher reif fürs Altpapier. Kein Zweifel: Das Higgs muss gefunden werden!

Wissenschaftler, die sich den kleinsten Bausteinen der Materie verschrieben haben, forschen mit gigantischen Anlagen. In Ermangelung anderer Beobachtungsgeräte bleibt ihnen nichts anderes übrig, als subatomare Teilchen stark zu beschleunigen und aufeinanderprallen zu lassen. Aus dem, was dann folgt, können sie Hinweise auf

die innere Struktur der beteiligten Teilchen erlangen.

Zum Betrieb geeignet sind alle frei vorkommenden, geladenen Atombausteine, also Elektronen und Protonen sowie deren spiegelsymmetrische Gegenstücke aus Antimaterie. Neutrale Teilchen sind ungeeignet: Sie reagieren nicht auf die elektrischen Leitfelder des Rings. Auch Quarks kann man nicht direkt verwenden, denn sie lagern sich infolge der starken Kernkraft sofort zusammen.

Doch auf der Suche nach dem Higgs-Teilchen schwiegen bislang alle Detektoren. Ganz aussichtslos ist die Sache nicht. Schon die Vetter des gesuchten Teilchens, die W- und Z-Bosonen – auch sie sind in die schwache Kernkraft involviert – gingen den CERN-Teams schon zu Zeiten des alten Beschleunigers ins Netz. Damals rasten noch die verhältnismäßig leichten Elektronen durch die Röhren. Als es keine Aussicht auf weitere Erkenntnisse über Higgs & Co gab, weil sich seine Leistung nicht weiter steigern ließ, wanderte der Elektronenbeschleuniger LEP (Large Electron and Positron Collider) ins Altmittel.

Mit fast Lichtgeschwindigkeit im Kreis

Begonnen hat alles 1931 mit einem Beschleuniger, der aussah wie eine Keksdose. Die nächste Generation fand in einer Garage Platz, dann folgten Fabrikhallen. Schließlich buddelten die Forscher ihre Beschleuniger unter die Erde. Die in Elektronenvolt gemessene Beschleunigungsenergie stieg von einigen Tausend Elektronenvolt in den Bereich Mega, Giga und heute Tera – Sprünge wie bei der Speicherkapazität der Festplatten. Entsprechend wuchs der Aufwand.

Noch in den 80er-Jahren gab es mehrere solcher Anlagen, alle von zumindest vergleichbarer Leistungsfähigkeit. Momentan existieren noch zwei Großprojekte: der 3,8 km lange Relativistic Heavy Ion Collider (RHIC) in Upton, New York, aus dem Jahr 2000 und der gut 6 km lange Tevatron in Batavia, Illinois, von 1987. Alle weiteren Großprojekte fielen dem Rotstift zum Opfer; denn jegliche Überlegungen, die gigantischen Kernkräfte in einer übermächtigen Quarkbombe zum Einsatz zu bringen, erwiesen sich bereits in den 90er-Jahren als nicht zielführend. Auch einem bereits ausgeschachteten Ring in Texas drehte der US-Senat den Geldhahn zu.

Spätestens, wenn der LHC seine volle Leistungsfähigkeit von 14 TeV erreicht, sollen die deutlich kleineren und langsameren Ringe Tevatron und RHIC in den Ruhestand gehen. Denn dann beschleunigt man in

dem 26,7 Kilometer langen LHC-Ring Protonen und Ionen auf das auf Erden machbare – auf 99,9999991 % der Lichtgeschwindigkeit. Dabei herrschen im Moment der Kollision Zustände wie am Anfang aller Zeiten, als sich – so glaubt man – kurz nach dem Urknall Energie und Materie voneinander schieden und sich die Naturgesetze herauskristallisierten.

Um die schweren Elementarteilchen in der Bahn zu halten, sind 9 Tesla starke Felder nötig. Das erfordert supraleitende, auf minus 268,7 Grad Celsius oder 4,3 Kelvin gekühlte Spulen, in denen der Strom widerstandslos zirkuliert. Dafür musste die Konstruktion des Beschleunigers, die Erdkrümmung und sogar die von den umliegenden Alpenmassiven verursachten Gravitationsasymmetrien berücksichtigen. In die Kalibrierung der Experimente gehen ebenso die Gezeiten und die Mondphasen ein.

Genau genommen besteht der LHC aus zwei Ringen, eng nebeneinander geschmiegt. In diesen beiden Röhren zirkulieren die Teilchen – momentan nur Protonen – gegenläufig. Beide Ringsysteme nutzen dieselben Magnetspulen und dasselbe Kühlsystem. Von außen sichtbar sind nur die blauen, jeweils 15 Meter langen Dipol-Röhren, die alles umschließen.

An vier Stellen des Rings eingelassen sind die Großexperimente – Detektoren – CMS, ATLAS und ALICE sowie der LHCb, die die eigentlichen Daten liefern. Im Fokus stehen neben dem Higgs-Teilchen das Quark-Gluon-Plasma, Verletzungen der CP-Invarianz (Charge Parity) sowie Mesonen, das Bottom-Quark und Neutrinos. Letztere entstehen bei den Experimenten in Massen. Ihre Analyse übernimmt ein italienisches Partnerinstitut in der Nähe von Rom.

– CMS: Der Compact Muon Solenoid zeichnet Proton-Proton-Kollisionen auf.

– ATLAS (A Toroidal LHC Apparatus): Er nimmt wie der CMS Proton-Proton-Kollisionen auf, ist aber anders konstruiert. Er dient unter anderem zur Kontrolle der beim CMS gewonnen Erkenntnisse.

– ALICE (A Large Ion Collider Experiment) ist ein Vielzweckdetektor, der den Zusammenprall schwerer Ionen aufzeichnen soll.

– LHCb: Das Experiment Large Hadron Collider beauty soll Symmetrieverletzungen und seltene Hadronen-Zerfälle beobachten.

Pleiten, Pech und Pannen

Doch das teuerste und wertvollste am LHC ist der Ring selbst. Wenn eines der vier großen Experimente nicht funktionieren würde, wäre das ärgerlich, aber keine Katastrophe. Der Ring aber mit seinen 1232

handgefertigten Dipolmagneten – ein jedes so teuer wie ein Einfamilienhaus – ist nicht zu ersetzen. Selbst eine Reparatur würde enorm viel Zeit verschlingen, da man die Röhren mühsam Stück für Stück aus 100 m Tiefe holen müsste.

Um so ärgerlicher für die Wissenschaftler, als sich bei Tests herausstellte, dass einer der zahlreichen Übergänge zwischen den Dipolen schadhaft war. Diese Übergänge sollen die Längenänderungen ausgleichen, die auftreten, wenn der Ring heruntergekühlt wird. Im Betrieb gehört der Ring zu den kältesten Orten im Universum. Denn nur in supraleitendem Zustand knapp über dem absoluten Nullpunkt lassen sich die Magnete energieeffizient betreiben. Um den Schaden zu lokalisieren, ließen die Techniker einen Tischtennisball mit Sender im Ring zirkulieren, der ihnen das defekte Element zeigte. Sonst hätten sie über Monate alle Verbindungen lösen müssen.

Weitere Schwierigkeiten bereiteten die sogenannten Quadropolmagnete, von denen jeweils drei zusammengehören. Sie pressen den Teilchenstrom kurz vor dem Kollisionspunkt auf ein Hundertstel Millimeter zusammen, um die Wahrscheinlichkeit einer Kollision zu erhöhen. Sie war bereits im Tunnel eingebaut, als sich im März 2007 ein Designfehler herausstellte – ihr Schutzmantel war zu schwach. Erst nach langem Nachdenken fanden die Ingenieure einen Weg, sie im Tunnel nachzubessern. Ein Umbau über Tage hätte Jahre gedauert. Eine derart lange Verzögerung hätte aber das Aus für das multinationale Projekt bedeuten können.

Dass die Quadropole aus den USA kamen, dem Land mit dem einzigen halbwegs vergleichbaren Beschleuniger Tevatron, der zudem mit dem Start des LHC seinem Ende entgegensieht, sorgte in langen Winter Nächten für Gesprächsstoff beim CERN.

Am 10. September 2008 folgte endlich die Eröffnung mit viel Prominenz. In den Folgetagen wurde die Energie im Ring zwar langsam gesteigert, aber wohl doch zu schnell. Eine der zahllosen elektrischen Übergänge erwies sich als kalte Lötstelle. Die enorme Stromstärke produzierte eine solche Hitze an dem schlecht leitenden Übergang, dass das Metall schmolz. Große Mengen des als Kühlflüssigkeit verwendeten Heliums traten mit lautem Knall aus und verschoben 53 weitere Magnete. Wieder schien kurze Zeit alles verloren. Immerhin dauerten die Reparaturarbeiten diesmal ein volles Jahr.

Seit Ende 2009 arbeiten die Magnete wieder. Am 30. März 2010 dann glückliche Gesichter. Zwar hatte man kein neues Teilchen zu verkünden, aber einen Weltrekord: Die Kerne pulsierten mit einer Energie von



Wenn der LHC seine volle Leistungsfähigkeit erreicht hat, fliegen Protonen und Ionen mit 99,9999991 % der Lichtgeschwindigkeit durch die Röhren (Abb. 2).

3,5 TeV durch den Genfer Untergrund. Das ergibt eine Kollisionsenergie von 7 TeV. Ein Jahr wird jetzt bei 7 TeV gemessen, allerdings zunächst mit wenigen Protonen, was die Aussicht auf spektakuläre Erfolge mindert. 2013, nach einem Umbau, soll es weitergehen. Das Ziel heißt dann: 14 TeV Kollisionsenergie und volle Teilchenzahl.

Wenn alles läuft

Unter Volllast erhoffen sich die Physiker dann mehr als 30 Millionen Kollisionen pro Sekunde. Obwohl Milliarden Atomkerne im Ring zirkulieren, ist die Wahrscheinlichkeit eines frontalen Zusammenstoßes zweier winziger Kerne extrem gering. Zudem legt der Zufall fest, mit welcher Energie und unter welchem Winkel die Protonen zusammenprallen. Nur hin und wieder entstehen dabei jene Teilchen, die es so seit dem Urknall nicht mehr gegeben hat, und auf die die Forscher so sehr gespannt sind. Sie möchten den ganzen Teilchenzoo vereinfachen und auf wenige Grundelemente reduzieren.

Die heute so komplizierten physikalischen Gesetze könnten damals ganz einfach gewesen sein. „In den ersten Momenten des Universums muss zwischen Teilchen und Kräften eine einfache Beziehung geherrscht haben, die uns heute verborgen ist“, vermutet der CERN-Physiker und Nobelpreisträger Carlo Rubbia. Diese verborgenen Beziehungen könnten dann endlich so sein, wie es der griechisch gebildete Abendländer am liebsten hat: klar und von einfacher, symmetrischer Struktur.

Doch das sind Spekulationen. Zunächst gilt es, nach jeder Kollision die darauf folgenden, rasend schnell ablaufenden Umwandlungen durch ganze Batterien von Sensoren zu erfassen, denn direkt lassen sie sich nicht beobachten. Noch in den 70er-Jahren leitete man die beim Zusammenprall entstehenden Teilchenschauer durch Kammern mit überhitzter Flüssigkeit und fotografierte die Spuren der Kernteilchen. Hier zeichneten sie ihre Flugbahn in Form von kleinen Bläschen in den instabilen Flüssigkeits-See, die dann fotografiert wurden.

Die Damen, die vor dreißig Jahren an schweren Pulten die Bilder digitalisieren mussten, sind nur noch Geschichte. Heute fangen Blei-Wolfram-Kristalle die Gammastrahlung der an den Prozessen beteiligten Teilchen auf und verwandeln sie proportional zu ihrer Energie in sichtbare Lichtblitze, die dann von Fotodioden in computergerechte, elektrische Signale umgewandelt werden. Allein der Detektor „ALICE“ besitzt 20 000 dieser Platten, zum Stückpreis von 200 Dollar. Ein Prinzip der Blaskammer ist geblieben: Nur die Folgeerscheinungen, also die Spuren der Ereignisse, lassen sich messtechnisch erfassen. Kein Wunder, dass die Physiker erst Monate oder gar Jahre nach einem Experiment genau wissen, was sie beobachtet haben. Ohne Rechenleistung läuft in diesem Feld der Physik nichts.

Geduld werden die Forscher auf jeden Fall brauchen. So zeigt sich das Higgs, wenn überhaupt, nur alle 10^8 Kollisionen. Wenn es soweit ist, fällt hoffentlich nicht der Rechner aus.

der Computing Grid (LCG), soll den Engpass beheben. Um den Datenfluss und Rechenaufwand organisatorisch in den Griff zu bekommen, haben seine Konstrukteure fünf Stufen oder Tiers vorgesehen: Das CERN-eigene Rechenzentrum bildet mit seinem Linux-Cluster und den großen Tape Libraries, die alle im RZ eintreffenden Daten horten, die Stufe 0 oder Tier 0.

Über eigene 10-Gbit-Leitungen fließen Kopien ins weltweite Speicher- und Rechen-Grid, in dem die eigentlichen Analysen stattfinden. Erste Stationen sind zwölf über den Erdball verstreute Rechenzentren mit großen Storage-Systemen und Clustern mit mehreren Tausend CPUs (Tier 1). Darüber verteilen sich die Daten und Rechenjobs auf über hundert Rechenzentren nationaler Institute und Universitäten (Tier 2) und schließlich auf die Server und Workstations der Arbeitsgruppen (Tier 3) und Anwender (Tier 4). Ab Mai wird das Eurogrid in Amsterdam die Organisation übernehmen.

Stufe 0 – der Genfer Linux-Cluster

Sah es lang so aus, als müsste die Physik auf die Informatik warten, war es letztlich umgekehrt. Durch zahlreiche Pannen und Fehler (siehe Kasten „Der Large Hadron Collider“) verzögerte sich der Start um mehr als drei Jahre.

Zeit genug für die IT, ihre Systeme zu aktualisieren.

Am meisten hat sich in den letzten Jahren am hauseigenen Linux-Cluster, dem Tier 0, getan. Dass seine Knoten alle drei Jahre dem Austausch zum Opfer fallen, führte dazu, das von der ursprünglichen IT des Jahres 2006, unmittelbar vor dem geplanten Start, heute außer den Kabeln und der Kühlung nichts mehr vorhanden ist. Die 5000 Server-Tower mit je zwei Single-Core-Xeons sind 6900 Mehrkern-Systemen mit insgesamt 41 000 Cores gewichen.

Lange versuchten die Vertreter von IBM, HP und anderen Computerherstellern vergeblich, die Leitung des Rechenzentrums von den Vorteilen von Blade-Rechnern zu überzeugen. Erst jetzt hatten sie Erfolg. Unterm Strich zählen für das CERN vor allem die Kosten und der in HEPSpec06 gemessenen Integer-Rechenleistung. Die Experimente sind schließlich schon teuer genug. „Wir sind Physiker“, erläuterte Wolfgang von Rüden, „Vorteile müssen sich rechnen.“ Außer der Rechenleistung zählt aufgrund der beschränkten Kühlkapazität vor allem die Power-Effizienz. Daher verwendet das Rechenzentrum momentan Quad-Sockel-Blades mit Vier-Kern-CPU's. Insgesamt steht im CERN eine Rechenleistung von 235 000 HS06 (= 60 Millionen SI2000) zur Verfügung – im gesamten Grid sind es garantierte 1 000 000 HS06.

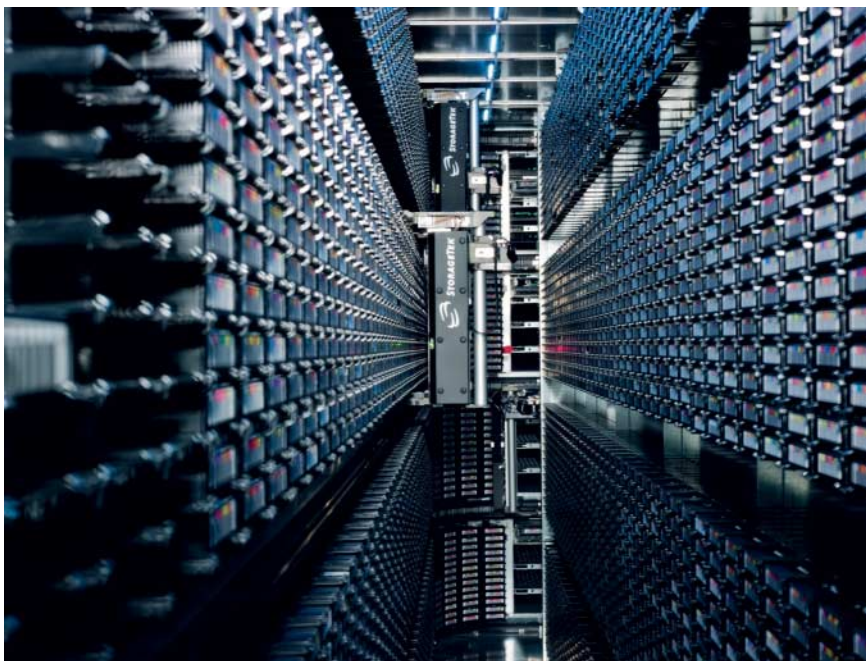
Raumangel herrscht im CERN-Rechenzentrum nicht, und die für die Administration nötigen Hilfskräfte sind auch in ausreichender Zahl vorhanden. Inzwischen ist auch das ganze CERN-Gelände mit redundant ausgelegten 10-Gigabit-Ethernet-Leitungen durchzogen. Gerne würde die IT-Abteilung noch mehr Cluster-Knoten ordern, doch wie so oft spielt die Kühlung nicht mit – das Rechenzentrum stammt aus den Mainframe-Tagen. Von den fünf Megawatt, die dem Rechenzentrum zur Verfügung stehen, wandern drei in die Rechner und zwei in die Kühlung. Doch wenn 2012 der Ring aufgerüstet ist, sind noch mehr Daten zu erwarten. Ab diesem Zeitpunkt, so die aktuellen Überlegungen, könnten kleine Rechenzentren, die in Containern untergebracht würden, das große ergänzen.

Momentan liefert der Ring in jeder Sekunde 2000 MByte Rohmaterial ans Rechenzentrum. Pro Jahr beläuft sich der Output des gesamten Rings auf über 20 PByte, das entspräche der Speicherkapazität von 4 Millionen DVDs. Als schneller, aber kurzfristiger Speicher dienen NAS-Server, deren Disk-Arrays 14 PByte auf 60 000 Disks fassen. Von dort verteilen sich die Aufnahmen der Ring-Sensoren zur Nachbearbeitung auf die Cluster-Knoten.

Der IT kommt entgegen, dass sich die dafür notwendigen Berechnungen leicht parallelisieren lassen – jeder CPU-Kern im Cluster bearbeitet jeweils ein Event. „Diese Berechnungen kann man sich vorstellen wie die Nachbearbeitung der Urlaubsbilder im RAW-Format auf dem heimischen PC“, erläutert Wolfgang von Rüden, „alle Bilder sind unabhängig voneinander. Es kommt darauf an, durch den Rechner Kontrast und Helligkeit zu optimieren, damit sich eine brillante Abbildung ergibt.“ Bei den Messdaten der CERN-Experimente sind es vor allem die örtlich und zeitlich variierenden Kalibrierungen der Sensoren, die der Rechner in die Rohdaten einarbeiten muss.

Viel Arbeit für die Tape-Roboter

Auf diese Daten müssen die mit der Analyse der Experimente befassten Physiker später jederzeit zugreifen können – die Archivdaten sind also Produktivdaten. Und damit die Theoretiker ihre Modelle auch präzise überprüfen können, ist jedes Bit gefragt.



Quelle: CERN

Insgesamt 48 000 Bänder können die 150 Enterprise-Bandlaufwerke in den Bibliotheken mit je 1 TByte Daten füllen (Abb. 3).

Das CERN – Europäische Organisation für Teilchenforschung

Gegründet wurde das CERN (Conseil Européen pour la Recherche Nucléaire) 1954 im dünnbesiedelten Grenzgebiet zwischen der Schweiz und Frankreich. Aus zunächst elf Mitgliedsstaaten wurden inzwischen zwanzig. 3000 Mitarbeiter sind beim CERN beschäftigt, vernetzt sind sie mit rund 8000 Gastwissenschaftlern von Universitäten und Forschungsinstitutionen. Hauptarbeitsgebiet des CERN sind Bau und Betrieb großer Teilchenbeschleuniger und der dazugehörenden Messtechnik. Das CERN erforscht weder Kernspaltung noch Kernfusion, sondern den inneren Aufbau von subatomaren Teilchen. Dazu beschleunigt man diese auf hohe Energien und lässt sie kollidieren. Nach dem Baustopp für eine vergleichbare Anlage in den USA hat das CERN in der Wissenschaft einen einzigartigen Status: Die durch CERN-Experimente gewonnenen Forschungsergebnisse lassen sich praktisch nur durch andere CERN-Experimente oder durch theoretische Berechnungen widerlegen. Die Physiker können nicht exakt steuern, welche Teilchen unter welchen Winkeln und Geschwindigkeiten kollidieren. Daher ist jede Kollision ein Unikat, sie lässt sich nicht gezielt wiederholen.

Schon früh erkannten die CERN-Wissenschaftler, dass es angesichts des streng limitierten Budgets nicht möglich sein würde, die gesamte Datenverarbeitung wie bei früheren Experimenten im eigenen Haus zu halten. Mehr Mittel könnte es nur geben, wenn dem alle Mitgliedsstaaten zustimmen würden. Aus der Not wurde eine Tugend. Ein genauso schneller wie sicherer Grid-Datenverbund schaufelt die Messwerte zu den angeschlossenen Partnerorganisationen. Das CERN behält aber die Übersicht. Es weiß stets, in welchem Knoten sich die Daten befinden. Mit Rechner-netzen kennt man sich am CERN aus. 1989 legte Tim Berners-Lee hier die Grundlage für das World Wide Web.

Auch unter Physikern ist das CERN umstritten gewesen. Zu teuer, zu schwerfällig und zu ineffizient, lautet der Vorwurf. Nur zwei Teilchen, das W- und das Z-Boson hat das CERN als Erstes entdeckt, und auch das nur, weil die eigentlich schnellere US-Gruppe ihre Ergebnisse so lange überprüfte, bis ihr das CERN um einen Tag den Ruhm und den Nobelpreis vor der Nase wegschnappte.

Die Befürworter sehen das anders. Ohne das CERN wäre das heute allgemein akzeptierte Standardmodell der subatomaren Teilchen, bestehend aus Quarks, Photonen, Mesonen und vielen anderen Partikeln, nicht so weitreichend bestätigt. Außerdem war das CERN der einzige Ort, an dem sich Spitzenforscher aus Ost und West auch während des Kalten Krieges ungestört treffen konnten. Ähnliches gilt auch heute noch für China und Taiwan.

Das CERN sucht ständig den Kontakt mit Universitäten und Fachhochschulen, um qualifizierte Gaststudenten zu bekommen. Es winken genauso spannende wie lehrreiche Monate in der Schweiz, denn alles, was es nicht in ausreichender Qualität von der Stange gibt, muss selbst programmiert werden. Eine Chance haben nicht nur Physiker, sondern auch Informatiker und Techniker. Linux-Fans dürften auf ihre Kosten kommen, denn das CERN setzt verstärkt auf dieses Betriebssystem. Generell ist Open Source gern gesehen. Bekannte IT-Firmen zahlen manchmal sogar Geld, um dem CERN bei seinen Spitzenleistungen helfen zu dürfen. Einen besseren Erfolgswachweis kann man sich schwer vorstellen. Dieses sogenannte Partnerprogramm ist ebenfalls eine Möglichkeit, eine Zeit lang als IT-Spezialist beim CERN Erfahrungen zu sammeln.

„Unsere Daten liefern ein holografisches Bild der Vorgänge in der Messkammer. Fehlen einige Datensätze, wird das Bild dadurch zwar nicht unvollständig, es nimmt aber immer mehr an Detailreichtum ab“, erläutert Bernd Panzer-Steindel, Physiker am CERN. Das bedeutet aber auch, dass alte Daten für neue Berechnungen überaus wichtig sind, und dass neue Erkenntnisse nur mit den Erfahrungen der Vergangenheit möglich sind. „Die alten Messwerte verbessern die Statistik“, ergänzt er, „gerade bei schwachen Effekten in der Nähe des statistischen Rauschens sind sie eine wichtige Hilfe, auf die wir nicht verzichten können.“

Deshalb bleibt – trotz weltweit verteilten Rechnens – eine Kopie aller ins Grid gestellter Daten beim CERN. Das sind immer noch 2000 MByte/s, die einerseits ins Tier 1 und gleichzeitig auf die eigenen Bandbibliotheken zu schieben sind. In ihnen füllen insgesamt 150 Enterprise-Bandlaufwerke des Typs TS1130 von IBM und StorageTek T10000B von Sun/Oracle parallel den momentan auf 48 PByte ausgelegten Langzeitspeicher, bestehend aus 48 000 Tapes mit je 1 TByte Kapazität. Hier lagern die Aufzeichnungen der laufenden Experimente einträchtig neben den Ergebnissen des alten Elektron/Positron-

Beschleunigerrings ebenso wie die berechneten LHC-Daten aus den Bau- und Test-Phasen. „Platten wären einfach zu teuer“, so Wolfgang von Rüden.

Da das Institut seinen gesamten Datenbestand alle vier Jahre auf die jeweils aktuelle Generation von Bandsystemen umkopiert, halten sich Ausfälle in Grenzen. Statistisch „lebt“ ein Datenträger zehn Jahre. Bei der großen Menge der eingesetzten Bänder macht das im Schnitt trotzdem ein defektes Band pro Monat. Können die CERN-Techniker dem lädierten Datenträger nicht aus eigener Kraft seine Geheimnisse entlocken, fliegt das widerspenstige Exemplar zu einem herstellereigenen Speziallabor in die USA.

2003 gelang es den Technikern des CERN im Test für den für Ende 2007 geplanten Start des Beschleunigers, mithilfe von 45 Stk-9940B-Bandlaufwerken über mehrere Stunden hinweg 1,1 GByte pro Sekunde auf Bänder zu schreiben, in Spitzen 1,2 GByte. Heute liegen die kontinuierlich speicherbaren Ströme im Mittel bei 2 GByte/s, maximal bei 4 GByte/s mit Spitzenraten von über 5 GByte/s.

Eine Kopie aller Bilder wandert nach Themen klassifiziert zur Analyse in die nächste Stufe des Grid, das sogenannte Tier 1. Die 12 dort zugehörigen

Rechenzentren stellen sowohl garantierte Rechenleistung als auch Speicherkapazität zur Verfügung. Jedes einzelne unterhält einen Rechen-Cluster mit jeweils mehreren Tausend CPUs sowie Disk- und Bandsystemen mit mehreren PByte Kapazität.

Stufe 1 bis 4 – zur Auswertung verteilt

Für die schnelle Weiterleitung der Datenfracht zwischen den Instituten sorgen die nationalen und internationalen Wissenschaftsnetze, ein separater Teil des Internet für die Vernetzung nichtkommerzieller Forschungseinrichtungen. Das CERN ist direkt mit dem im Juni 2005 in Betrieb genommenen Europäischen Backbone GÉANT2 verbunden, der aus Dark Fiber (nicht an bestimmte Protokolle gebundene Glasfaser-Bündel) und gemieteten Leitungen besteht. Über Erstere lassen sich per Wellenlängen-Multiplexing 10 GBit/s pro Wellenlänge und Faser übertragen.

Vom GÉANT gehen die Daten zum einen über die daran angeschlossenen nationalen Netze (National Research and Education Networks, NREN) an die beteiligten europäischen Tier-1-Knoten CNAF (Bologna), IN2P3 (Lyon),

GridKa (Karlsruhe), SARA (Amsterdam), NorduGrid (Skandinavien), PIC (Barcelona) und RAL (Didcot, UK). Über den großen interkontinentalen Backbone gelangen sie zum anderen durch das kanadische CA*net4 nach Vancouver, BC (TRIUMF) sowie in das US-amerikanische ESnet, von hier aus zu den Knoten Brookhaven (Upton, NY), Fermilab (Batavia, Illinois) sowie über das ASnet zum ASCC in Taipeh. Für die gesamte Tier-0-Tier-1-Vernetzung sind in allen beteiligten Netzen mehrere 10-Gbit-Links permanent reserviert und bilden das Optical Private Network (OPN) des LCG.

Etwa 100 Universitätsrechenzentren in 40 Ländern bilden das Tier 2. Sie stellen weitere garantierte Rechenkapazität und sogenannte Managed Disk Storage – hauptsächlich für Simulationen und Analysen – bereit. Außerdem gewähren sie den LCG-Zentren des Tier 3 Zugang zum Grid und stellen ausreichend Netzwerkbandbreite und entsprechende Dienste für den Datenaustausch mit den Tier-1-Knoten bereit.

Weitere 1000 Institute, die Berechnungen mit den CERN-Daten vornehmen, stellen das Tier 3 dar, die lokalen Arbeitsplatzrechner von etwa 8000 Wissenschaftlern das Tier 4. Denn je nachdem, was bei der betreffenden Kollision passiert ist, interessieren sich unterschiedliche Wissenschaftler für das Ereignis. Die betreffenden Gruppen können sich die Daten kopieren und mit ihren Formeln weiterbearbeiten. Oft reicht dazu ein Desktop-PC.

Sind die jeweiligen Berechnungen im Tier 2 und darunter abgeschlossen, geben sie die ursprünglichen und die

errechneten respektive durch Simulationen erzeugten Datensätze an das zuständige Tier-1-Zentrum zurück, denn nur hier stehen die Bandspeicher für die Langzeitspeicherung. Das CERN kann bei Bedarf ebenfalls auf diese Daten zugreifen. Die Experimentdaten werden aus Sicherheitsgründen mehrfach, in unterschiedlichen Tier-1- und Tier-2-Zentren vorgehalten.

Für die Verteilung der Daten an die deutschen Tier-2-Zentren ist das GridKa (Grid am Forschungszentrum Karlsruhe) zuständig. Die Anbindung sowohl ans GÉANT als auch an die Tier-2-Zentren sowie deren Anbindung an die Tier-3-Standorte übernimmt das deutsche Forschungsnetz X-WIN, das Anfang 2006 das Wissenschaftsnetz G-WIN ablöste.

Karlsruhe selbst ist momentan über eine redundante 20 Gbit/s schnelle Leitung an das Grid angeschlossen. Das Forschungszentrum stellt eine Rechenkapazität von knapp 22 250 kSI2k sowie über 1 PByte auf Disks und 15 PByte auf LTO-3- und LTO-4-Bändern zur Verfügung. Das deutsche GridKa erhielt am 20.11.2009, unmittelbar nach dem erneuten Start des Rings, als weltweit erstes Schicht-1-Zentrum die Daten des Experiments CMS zur Bearbeitung. Kurze Zeit später trafen auch die Daten der anderen LHC-Experimente in Karlsruhe ein.

Auf europäischer Ebene verwaltete das EGEE (Enabling Grids for E-science) die Rechenressourcen für wissenschaftliche Anwendungen und damit auch für den LHC. Allerdings endete EGEE am April 2010. Die neue European Grid Initiative (EGI) mit Sitz in

Amsterdam soll dessen Rolle übernehmen. Sie wird für zunächst vier Jahre gefördert. Beteiligt sind etwa 34 nationale Grid-Initiativen in Europa. Eine Aufgabe von EGI wird es sein, die technisch verwandten Bereiche Grid und das kommerziell geprägte Thema Cloud zu verschmelzen. Nicht nur der LHC, sondern auch datenintensive Forschungsbereiche wie die Systembiologie sollen vom Grid/Cloud in Europa profitieren. Dazu sollen die Cloud-Ressourcen über das Grid zugänglich gemacht werden. Vier Firmen und zwei Universitäten haben sich hierzu im Projekt „Open Cirrus“ zusammengeschlossen. Die offene Wolke soll den Sprung vom Grid zum Cloud erleichtern.

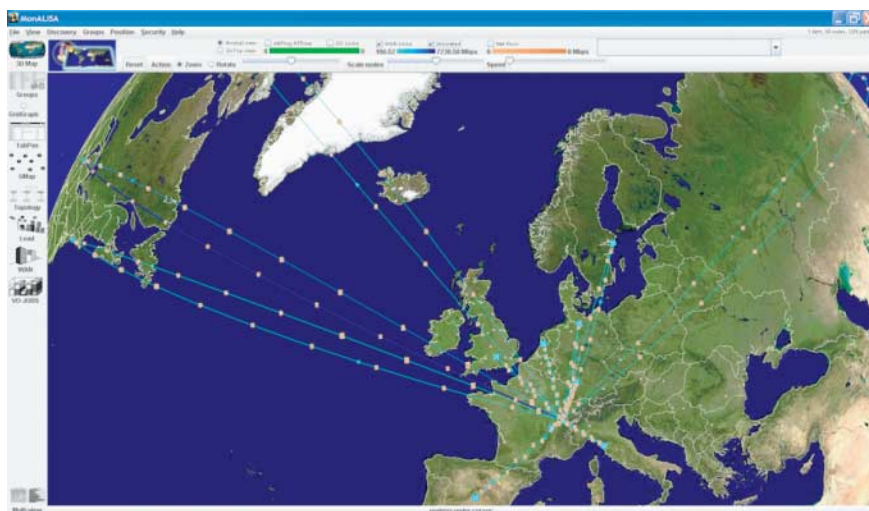
Egal, wie es mit den Fördergeldern der EU weitergeht, bereits im Vorfeld des Übergangs von EGEE zu EGI einigten sich die wichtigsten CERN-Staaten darauf, das lebensnotwendige Rechnernetz für den Beschleunigerring auf jeden Fall weiter zu unterhalten. Budgetbedingte Stellenkürzungen beim CERN schließt das allerdings nicht aus.

LHC Computing für zu Hause

Neben dem LCG setzt das CERN noch auf ein weiteres Rechnernetz, nämlich das seiner Freunde. Die stellen, ähnlich wie beim SETI-Projekt, die Rechenpower ihrer unterbeschäftigten PCs zur Verfügung. Das CERN spricht diese Rechner über das normale Internet an und bündelt sie zu einer schlagkräftigen Armada. Diese zusätzlichen Rechner schultern vor allem lästige, aber nicht unbedingt zeitkritische Simulationsberechnungen der Großexperimente. Obwohl die CERN-Wissenschaftler ihr virtuelles Hilfsrechenzentrum ohne großen Werbefeldzug in Umlauf brachten, überstieg die Rechenkapazität an Spitzentagen schon rasch die des eigenen Rechenzentrums.

Jeder der will, kann bei LHC@home auch weiterhin mitmachen und wird mit dem Gefühl belohnt, am größten Experiment der Welt teilzunehmen (siehe Kasten „Links zu CERN-Projekten“). Die Kapazität der freiwilligen Helfer ist beim CERN fix eingeplant. Von Afghanistan bis Simbabwe ist fast jedes Land der Erde vertreten. Viele der freiwilligen Helfer sind in den derzeit fast 5400 registrierten Teams organisiert, von denen das größte über 2200 Mitglieder zählt.

Neben einem Sicherheits-Backup der Rohdaten bleibt das zentrale Register



Eine Kopie der im LHC anfallenden und aufbereiteten Sensor-Daten wandern ins LHC Computing Grid und damit über den gesamten Globus (Abb. 4).

aller Messdaten beim CERN. Das Institut weiß also stets, welche Partnerorganisation welche Bits auf welchen Tape- oder Disk-Systemen in Verwahrung hat und führt zudem Buch über mögliche Versionsabweichungen der Daten im Grid von denen im eigenen Archiv. „Dazu haben wir eine spezielle Middleware entwickelt, sie hält uns Versionskollisionen vom Leib“, so von Rüden.

Kein Grid ohne Software

Da die beteiligten Institute unmöglich alle immer auf dem gleichen Softwarestand zu halten sind, soll das Grid genügend Intelligenz mitbekommen, um jedem Teilnehmer die Informationen „mundgerecht“, seinem Softwarestand entsprechend, zu liefern. Dies übernimmt ein Release-Management für die Middleware.

Fünf Mitarbeiter feilen außerdem ständig an den Sicherheitsmechanismen des Grid, um Angriffe über das Netz abzuwehren; weltweit sind rund 50 Personen ständig damit befasst. Zwar haben die CERN-Experimente weder einen militärischen noch einen kommerziellen Nutzen, doch jeder Zusammenprall ist ein Unikat, das man ungern verlieren möchte. Außerdem sollen sie verhindern, dass Cracker das Rechen-Grid kapern und es als mächtige Waffe gegen Server für DoS-Angriffe einsetzen. Schon jetzt zählt das Rechenzentrum allein auf den Tier-0-Cluster mindestens zwei Kaper-Versuche täglich.

Wie bei der voluminösen Hardware gilt auch bei der Software die CERN-Philosophie: Alles, was bereits vorhanden ist, wird benutzt, der Rest selbst entwickelt. Um das Budget zu schonen, setzt man bevorzugt auf freie Software oder, wo die nicht weiterhilft, auf Partnerschaften mit kommerziellen Herstellern. Erst danach wird dazugekauft – oder selber programmiert. So sind wesentliche Teile der Speichersoftware im CERN entstanden. „Die verfügbare Software war für unsere Anforderungen einfach nicht schnell genug“, so von Rüden. „Unser CASTOR (CERN Advanced Storage Manager) sorgt etwa dafür, das die Nutzer statt diverser Bandsilos nur einen einzigen virtuellen Datenträger zu sehen bekommen, den Rest regelt das Programm im Hintergrund.“

Die selbstentwickelte Software stellen die Grid-Entwickler ihrerseits wie-

Links zu CERN-Projekten

LCG	lcg.web.cern.ch/LCG
EGEE	public.eu-egee.org
EGEE-Lizenz	public.eu-egee.org/license/license2.html
LHC@home	lhathome.cern.ch
Castor	castor.web.cern.ch/castor
LHC	lhc.web.cern.ch/lhc
CMS	cmsinfo.cern.ch/Welcome.html
ATLAS	www.atlas.ch
ALICE	aliceinfo.cern.ch
LHCb	lhcb.web.cern.ch/lhcb

der für andere Projekte zur Verfügung. Eine Open-Source-Lizenz des EGEE regelt die Bedingungen für Umgang mit und Weitergabe der Grid-Software wie etwa der Lightweight Middleware for Grid Computing (gLite).

Als Plattform für eine Partnerschaft mit IT-Firmen hat sich das CERN das „Open Lab“ erdacht. Und da sich mit Namen des renommierten Forschungszentrums gut werben lässt, zahlen Konzerne minimal 800 000 Schweizer Franken (500 000 Dollar), um im Open Lab mitarbeiten zu dürfen. Hier testen sie Dinge, die erst Jahre später auf den Markt kommen.

Fazit

Endlich läuft der Ring und mit ihm die Experimente, wenn auch nur mit halber Kraft – mit $2 \times 3,5$ statt 2×7 TeV. Doch auch das ist deutlich mehr als irgendwo sonst auf der Welt und dürfte für neue Erkenntnisse locker reichen. Denn bereits ab 3 TeV erwarten die Physiker ein supersymmetrisches Teilchen, das es seit dem Urknall nicht mehr gegeben hat.

Auch wenn noch keine weltbewegenden physikalischen Entdeckungen durch den LHC vorliegen – eines hat das CERN bereits erreicht: An dem Ort, an dem das World Wide Web entstanden ist, haben sich nun alle renommierten Forschungseinrichtungen dieser Welt virtuell zusammengefunden – wieder über ein Netz, das LHC Computing Grid. (sun)

BERND SCHÖNE

ist freier Journalist.

 www.ix.de/ix1014058



Anzeige